



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>4</sup> : C12P 21/00, C12N 15/00 C07H 21/04		A1	(11) International Publication Number: WO 87/00865 (43) International Publication Date: 12 February 1987 (12.02.87)
(21) International Application Number: PCT/US86/01599 (22) International Filing Date: 29 July 1986 (29.07.86)		(74) Agent: ROWLAND, Bertram, L.; Leydig, Voit & Mayer, 350 Cambridge Avenue, Suite 200, Palo Alto, CA 94306 (US).	
(31) Priority Application Number: 760,236 (32) Priority Date: 29 July 1985 (29.07.85) (33) Priority Country: US		(81) Designated States: AT (European patent), AU, BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).	
(71) Applicant: CALGENE, INC. [US/US]: 1920 Fifth Street, Suite F, Davis, CA 95616 (US). (72) Inventors: GOODMAN, Robert, M. : 4037 Vista Way, Davis, CA 95616 (US). KNAUF, Vic, C. : 2454 Elen-dil Lane, Davis, CA 95616 (US). HOUCK, Catherine, M. : 172 Carmel Court, Vacaville, CA 95688 (US). CONIAL, Luca : 1892 Sharon Avenue, Davis, CA 95616 (US).		Published With international search report.	

(54) Title: MOLECULAR FARMING

## (57) Abstract

Novel constructs for expression of physiologically active mammalian proteins in plant cells, either in culture or under cultivation. The constructs provide a promoter functional in a plant host, a structural gene coding for mammalian protein and a terminator functional in a plant host. The construct is introduced into a plant cell to become integrated into the plant genome for expression in the plant cells or plants. The plant cells may be harvested and the mammalian protein isolated in physiologically active form.

## MOLECULAR FARMING

BACKGROUND OF THE INVENTION5 Field of the Invention

10 An extensive literature has been developed demonstrating the ability to produce polypeptide sequences in a wide variety of cellular hosts. Numerous genes have been isolated from mammals and viruses, joined to transcriptional and translational initiation and termination regulatory signals from a source other than the structural gene and introduced into hosts in which the regulatory signals are functional. Frequently, the peptide of interest is prepared to be used for a physiological purpose. In many situations, physiological activity requires not only having the correct or substantially correct amino acid sequence, but the peptide must fold properly including proper disulfide linkage formation and may require further processing, such as acetylation, glycosylation or methylation.

20 For economic production, one would wish to use unicellular microorganisms, which could be grown in large fermentation tanks, do not have fastidious nutrient requirements and are relatively economical to maintain. Bacteria, such as E. coli, B. subtilis, or the like, fungi, such as yeast, Candida, filamentous fungi, or the like, offer economic opportunities to produce a wide variety of peptides. However, because of the substantial difference in the nature of the unicellular microorganisms and mammalian cells, the folding and processing in a mammalian cell appears to be substantially different from these lower order organisms. Therefore, the products which are obtained from the unicellular microorganisms may not have been properly processed or folded so as to realize a substantial proportion or all of the physiological

35

activity of the naturally occurring peptide obtained from a native host.

There therefore remains substantial interest in providing alternative economic systems for producing peptides, where high yields may be obtained and significantly, the products may be produced in a form providing for a high degree of physiological activity common to the wild-type peptide having the same or substantially the same amino acid sequence.

#### Brief Description of the Relevant Literature

References concerned with expression of various interferons include Goeddel et al., Nucleic Acids Res. (1980) 8:4057-4074; Goeddel, Nature (1980) 287:411-415; Yelverton et al., Nucleic Acids Res. (1981) 9:731-741; Gray et al., Nature (1982) 295:503-508; Devos et al., Nucleic Acids Res. (1982) 10:2487-2501; Grey and Goeddel, Proc. Natl. Acad. Sci. USA (1983) 80:5842-5846; Scahill et al., Proc. Natl. Acad. Sci. USA (1983) 80:4654-4658. See also, Horsch et al., Science (1985) 227:1229-1231.

Wide host range cloning vectors are described by Knauf and Nester, Plasmid (1982) 8:45-54. The nucleotide sequence of the T-DNA region is described by Barker et al., Plant Molecular Biology (1983) 2:335-350. See also, EPA 0 116 718 and PCT WO84/02913.

#### SUMMARY OF THE INVENTION

Efficient production of physiologically active mammalian proteins is provided by introducing functional constructs containing the mammalian structural gene into a plant cell. The construct is able to express the desired peptide in an isolatable form. The plant cells may be grown in culture or cultivated in an appropriate nutrient medium or soil and the mammalian protein harvested. Particularly, T-DNA transformation may be employed for integration of

the construct into the plant genome under conditions where the cells can be used to produce plants.

#### DESCRIPTION OF THE SPECIFIC EMBODIMENTS

5 Novel DNA sequences are provided comprising constructs functional in plant cells for the expression of physiologically active mammalian peptides. The constructs provide for functional transcriptional and translational initiation and termination regulatory  
10 signals which with the construct integrated into the plant genome, provide for efficient expression of the structural gene. Depending upon whether the cells are grown in culture or cultivated in soil or other medium, these cells or plants may be harvested, and the desired  
15 product extracted and purified in accordance with conventional techniques.

The construct for introduction into the plant will be considered first. While the construct may be either DNA or RNA, for the most part, the construct  
20 will be DNA even though the construct may be ultimately introduced into the plant cell as RNA, such as a virus. One significant element of the construct will be the transcriptional cassette involving the transcriptional initiation region, which may be divided into a  
25 regulatory domain and the usually proximal RNA polymerase binding domain, the structural gene coding for the mammalian peptide of interest, and a terminator region which provides for termination of transcription and translation. Another significant element is the  
30 regulation of expression at a particular stage of differentiation, in a particular plant part, e.g., roots, leaves, stalk, or the like. In many situations a polypeptide leader signal will be desirable which directs the product to a particular organelle. This  
35 will be of substantial significance where the organelle may be involved in the proper processing of the peptide.

A wide variety of transcriptional initiation regions may be employed, which are inducible, e.g., tissue specific, or constitutive. The transcriptional initiation regions may come from the tumor-inducing plasmids of Agrobacterium, particularly the genes associated with T-DNA, or from viruses, or plants. Among the T-DNA transcription initiation regions which may find use include those regions associated with octopine synthase, nopaline synthase, mannopine synthase, transcript 7, and the like. Among viral transcription initiation regions are included the caulimovirus full length promoter and the region VI promoter. Among plant transcription initiation regions are the ribulose-1,5-bisphosphate carboxylase small subunit region, which is light inducible, or the napin or other seed protein region, for formation in seed.

The transcription initiation regions may be isolated from their natural sites in a host or may be sequenced and synthesized so as to have the same or substantially the same sequence as the wild-type region. Where inducible regulation is desired, domains may be obtained from different sources, so that a regulatory domain may be obtained from one source and joined to an RNA polymerase binding domain from a different source. In this manner, one may provide for the use of strong RNA polymerase binding domain associated with one structural gene, while having a regulatory domain associated with a different structural gene. These hybrid regulatory regions may find particular use where one wishes to induce the production of the desired gene at a particular phase in the growth of the plant or to have the product located in a particular plant part, such as seed, leaves, or the like. Any transcriptional termination regulatory region may be used which is functional in plants, e.g., prokaryotic or eukaryotic, from T-DNA, plants, viruses or mammals.

The structural gene may be any mammalian gene of interest, which includes mammalian viral pathogen genes. A wide variety of genes have been isolated and shown to be capable of production in unicellular micro-organisms, to various degrees of biological activity and efficiencies, and in mammalian cells, with the ever present concern that the mammalian cells are transformed cells, so that any product must be carefully purified to ensure the complete absence of any nucleic acid contaminant. Structural genes of interest include  $\alpha$ -,  $\beta$ - and  $\gamma$ -interferons, immunoglobulins, with the structural genes coding for the light and heavy chains and desirably assembly occurring in the plant cell, lymphokines, such as interleukins 1, 2 and 3, growth factors, including insulin-like growth factor, epidermal growth factor, platelet derived growth factor, transforming growth factor- $\alpha$ , - $\beta$ , etc., growth hormone, insulin, collagen plasminogen activator, blood factors, such as factors I to XII, histocompatibility antigens, enzymes, or other mammalian proteins, particularly human proteins.

Among the antigens associated with viral pathogens, include the core and envelope proteins of leukemia and lymphotropic retroviruses, such as HTLV-I, -II and -III, feline leukemia virus, etc., surface antigens of herpes simplex virus, hepatitis B virus, adenovirus, and the like.

Peptides of interest other than those indicated above will be peptides which may be administered physiologically, where growth in plants diminishes the probability of contaminants causing an adverse response upon administration to a mammalian host.

Plants may also find use in preparing other proteins than to be administered to a host, where the mammalian protein, such as an enzyme, may require folding and/or processing that is unavailable in unicellular microorganisms. Not only may plants be used to prepare the mature peptide, but in many

instances it may be desirable to prepare the precursor, which may require cleavage and assembly, either endogenous or exogenous to the plant. Peptides can be prepared having a naturally occurring transit or leader peptide or a transit or leader peptide from a plant peptide. The transit peptide will serve to subject the entire peptide product to the processing and maturing of the peptide. Such processing may include specific peptide cleavage, e.g., removal of the transit peptide, glycosylation at glycosylation sites, folding with appropriate formation of disulfide linkages.

Any convenient terminator may be employed which provides for efficient termination in conjunction with a particular transcriptional initiation region. The terminator may be the terminator region normally associated with the transcriptional initiation region or associated with a different transcriptional initiation region, so long as the region is functional in plants.

In order to select for plant cells that have successfully integrated the construct, the expression construct or cassette will usually be joined to a marker. The marker will allow for either screening or selection, usually selection. A number of different antibiotic resistance genes are known that can find use in plants to serve as markers. These genes include enzymes providing resistance to kanamycin, chloramphenicol, G418, gentamycin, and the like. These genes will have transcriptional and translational initiation and termination regulatory regions that may be the same or different from the regions employed for the structural gene of interest. Usually constitutive expression will be provided, rather than inducible expression.

In addition to the cassette and marker, depending upon the manner in which the DNA construct will be introduced into the plant cell, other sequences may be necessary. Where the Agrobacterium tumor-

inducing system is to be employed, one or both of the T-DNA boundaries of a Ti- or Ri-plasmid will be present, particularly the right boundary region. Each boundary region will generally be of about 1 to 1.5kbp.

5           The DNA construct of the cassette, marker and T-DNA may then be employed in a variety of ways. For integration into a Ti- or Ri-plasmid, the construct may be introduced into an appropriate Agrobacterium strain carrying the tumor-inducing plasmid, whereby the  
10       construct will become integrated into the tumor-inducing plasmid and may then be transferred to plant cells. Alternatively, the construct may be joined to a broad spectrum replication system, such as a P1 incompatibility  
15       replication system and transformed into an Agrobacterium containing an armed or disarmed tumor-inducing plasmid. Integration will occur and transfer to the plant cell of the construct along with other genes and markers. When transfer is with an armed tumor-inducing plasmid  
20       (Ti or Ri T-DNA containing plasmid) genes conferring tumor formation will be transferred, so that galls may form. With disarmed tumor plasmids (lacking T-DNA) the tumor-causing genes cannot be transferred and gall formation is not encountered. A transposon may be  
25       employed containing the construct and the gene coding for transposase, such as in the Ac-Ds system. A viral system may be employed which provides for integration into the host genome.

          Transfer of the DNA construct into the plant cell may be by infection with A. tumefaciens or A.  
30       rhizogenes, microinjection, liposome fusion, viral infection, or the like. The particular manner in which the DNA is introduced into the plant cell for integration is not critical to this invention.

          Usually, the construct will be joined to a  
35       prokaryotic replication system, for example, a system functional in E. coli, so as to allow for cloning at the various stages of preparation of the construct. A



wide variety of replication systems are available, both plasmid and phage, such as ColE1,  $\lambda$ , etc.

Plant cells which are employed may be either monocots or dicots and will be chosen in accordance with the manner in which the desired gene is to be produced and harvested. Plants which may find use include tobacco, sunflower, corn, sugar cane, soybean, tomato, alfalfa, mustard, sugar beet, rapeseed, etc. The product may be found in plant parts such as seed, leaves, fruit, roots, stalks, tubers, or combinations thereof. Thus, the peptide of interest may be the sole purpose for growing the plant or be an additional product.

The construct will be prepared in conventional ways. DNA sequences may be detected by employing probes, which may be designed based on known amino acid sequences or prior isolation of all or fragments of mRNA or chromosomal DNA. The sequences may be restriction mapped or sequenced and the entire gene obtained by various techniques, such as walking, using a plurality of primers, or the like. Once the sequence has been isolated, it may be ligated to other sequences, either directly or through linkers, where the linkers may provide no or portions of coding sequences. Various strategies may be devised based on available restriction sites, the absence of restriction sites, the ability to introduce restriction sites, the availability of particular fragments, the presence of sequences which require excision, and the like. The particular strategy will be dependent upon the gene which is employed, the particular regulatory systems, the availability of vectors having one or more of the desired sequences, as well as other practical considerations.

The manner in which the construct is introduced into plants may be varied widely. This has already been indicated by virtue of the different

sequences which may be included in the construct. Of particular interest is the presence of T-DNA for integration in conjunction with the vir genes of the Ti-plasmid which may be present on a plasmid other than the plasmid containing the foreign gene. Thus, the necessary genetic capability for integration into the plant cell can be provided, in conjunction with infection with Agrobacterium or introduction of the DNA by other means. Descriptions of introduction of DNA into plants may be found in Pedersen et al., Plant Cell Reports (1983) 2:201-204; Hooykaas-Van Slogteren et al., Nature (1984) 311:763-764; de Cleene and de Ley, The Botanical Review (1976) 42:389-466, and references cited therein, are incorporated herein by reference.

The transformed plant cells will then be grown in appropriate nutrient medium to provide for selected calli, where plant cells or protoplasts have been modified. Once the calli has formed, the medium may then be changed to encourage root and shoot formation and the resulting shoots transferred to appropriate growth medium for growth of plants. When the plants have been grown to the desired stage, the plants or plant parts, e.g., seeds, fruit or the like, may be harvested, and the desired product isolated in accordance with conventional ways. Thereafter, the gene may be regenerated from seeds, so that the process of regeneration from calli need not be repeated. The plant may be ground and extracted with appropriate solvents, chromatographed, crystallized, solvent extracted, etc. The crude product may then be purified in accordance with the nature of the product.

In some instances it may be neither necessary nor desirable to extract and isolate the mammalian protein product from the plant. Where the product can have a physiological effect on ingestion, it may be sufficient that the product be retained with the plant. This will be true where the plant part is edible, such

as fodder which could include nutritional qualities, such as bovine growth hormone, seed, nuts, fruit, and vegetables, which could include proteins involved in the regulation of digestion, or the like.

5 The following examples are offered by way of illustration and not by way of limitation.

#### EXPERIMENTAL

10 The HindIII-SmaI fragment of Tn5 containing the entire structural gene for APHII (Jorgensen et al., Mol. Gen. (1979) 177:65) was cloned into pUC8 (Vieira and Messing, Gene (1982) 19:259), converting the fragment into a HindIII-EcoRI fragment, since there is an EcoRI site immediately adjacent to the SmaI site.

15 The PstI-EcoRI fragment containing the 3'-portion of the APHII gene was then combined with an EcoRI-BamHI-SalI-PstI linker into the EcoRI site of pUC7 (pCGN546W). Since this construct does not confer kanamycin resistance, kanamycin resistance was obtained

20 by inserting the BglII-PstI fragment of the APHII gene into the BamHI-PstI site (pCGN546X). This procedure reassembles the APHII gene, so that EcoRI sites flank the gene. An ATG codon was upstream from and out of reading frame with the ATG initiation codon of APHII.

25 The undesired ATG was avoided by inserting a Sau3A-PstI fragment from the 5'-end of APHII, which fragment lacks the superfluous ATG, into the BamHI-PstI site of pCGN546W to provide plasmid pCGN550.

The EcoRI fragment containing the APHII gene

30 was then cloned into the unique EcoRI site of pCGN451, which contains an octopine synthase cassette for expression, to provide pCGN552.

pCGN451 includes an octopine cassette which contains about 1556bp of the 5' non-coding region fused

35 via an EcoRI linker to the 3' non-coding region of the octopine synthase gene of pTiA6. The pTi coordinates are 11,207 to 12,823 for the 3' region and 13,643 to

15,208 for the 5' region as defined by Barker et al.,  
Plant Mol. Biol. (1983) 2:325.

The 5' fragment was obtained as follows. A  
small subcloned fragment containing the 5' end of the  
5 coding region, as a BamHI-EcoRI fragment was cloned in  
pBR322 as plasmid pCGN407. The BamHI-EcoRI fragment  
has an XmnI site in the coding region, while pBR322 has  
two XmnI sites. pCGN407 was digested with XmnI,  
resected with Bal31 nuclease and EcoRI linkers added to  
10 the fragments. After EcoRI and BamHI digestion, the  
fragments were size fractionated, the fractions cloned  
and sequenced. In one case, the entire coding region  
and 10bp of the 5' non-translated sequences had been  
removed leaving the 5' non-transcribed region, the mRNA  
15 cap site and 16bp of the 5' non-translated region (to a  
BamHI site) intact. This small fragment was obtained  
by size fractionation on a 7% acrylamide gel and  
fragments approximately 130bp long eluted. This size  
fractionated DNA was ligated into M13mp9 and several  
20 clones sequenced and the sequence compared to the known  
sequence of the octopine synthase gene. The M13  
construct was designated pl4, which plasmid was  
digested with BamHI and EcoRI to provide the small  
fragment which was ligated to a XhoI to BamHI fragment  
25 containing upstream 5' sequences from pTiA6 (Garfinkel  
and Nester, J. Bacteriol. (1980) 144:732) and to an  
EcoRI to XhoI fragment containing the 3' sequences.  
The resulting XhoI fragment was cloned into the XhoI  
site of a pUC8 derivative, designated pCGN426. This  
30 plasmid differs from pUC8 by having the sole EcoRI site  
filled in with DNA polymerase I, and having lost the  
PstI and HindIII site by nuclease contamination of  
HincII restriction endonuclease, when a XhoI linker was  
inserted into the unique HincII site of pUC8. The  
35 resulting plasmid pCGN451 has a single EcoRI site for  
the insertion of protein coding sequences between the  
5' non-coding region (which contains 1,550bp of 5'

non-transcribed sequence including the right border of the T-DNA, the mRNA cap site and 16bp of 5' non-translated sequence) and the 3' region (which contains 267bp of the coding region, the stop codon, 196bp of 3' non-translated DNA, the polyA site and 1,153bp of 3' non-transcribed sequence). pCGN451 also provides the right T-DNA border.

The resulting plasmid pCGN451 having the ocs 5' and the ocs 3' in the proper orientation was digested with EcoRI and the EcoRI fragment from pCGN551 containing the intact kanamycin resistance gene inserted into the EcoRI site to provide pCGN 552 having the kanamycin resistance gene in the proper orientation.

This ocs/KAN gene was used to provide a selectable marker for the trans type binary vector pCGN587.

The 5' portion of the engineered octopine synthase promoter cassette consists of Tia6 DNA from the XhoI at bp 15208-13644 (Barker's numbering), which also contains the T-DNA boundary sequence (border) implicated in T-DNA transfer. In the plasmid pCGN587 the ocs/KAN gene from pCGN552 provides a selectable marker as well the right border. The left boundary region was recloned from the HindIII-EcoRI fragment as a KpnI-EcoRI fragment in pCGN565 to provide pCGN580. pCGN565 is a cloning vector based on pUC8-Cm, but containing pUC18 linkers. pCGN580 was linearized with BamHI and used to replace the smaller BglII fragment of pVCK102 (Knauf and Nester, Plasmid (1982) 8:45), creating pCGN585. By replacing the smaller SalI fragment of pCGN585 with the XhoI fragment from pCGN552 containing the ocs/KAN gene, pCGN587 was obtained.

pCGN807 is derived from pBR322, contains the tetracycline resistance gene, has about 300bp of a sequence including the trp promoter 5' to a 750bp sequence including all of the murine  $\gamma$ -interferon

structural cDNA and a 3'-untranslated region, with the gene described in Gray and Goeddel, Proc. Natl. Acad. Sci. USA (1983) 80:5842-5846. The promoter and gene are an EcoRI fragment of about 834bp. pCGN807 was  
5 digested with EcoRI and HpaI which fragment was inserted into pEMBL18+, which had been digested with EcoRI and SmaI creating pCGN808. pEMBL18+ (Dente et al., Nucleic Acids Res. (1983) 11:1645) differs from the pEMBL plasmids in having the pUC18 linker inserted  
10 into the  $\beta$ -galactosidase gene. pCGN808 was digested with EcoRI and BamHI to provide a 540bp fragment which was inserted into pCGN46 which had been digested with EcoRI and BamHI to provide for the murine  $\gamma$ -interferon cDNA gene to be in the proper orientation in relation  
15 to the mannopine synthase (mas) 5' transcription initiation region and the ocs 3' transcription termination region (pCGN809).

An approximately 5.5kbp EcoRI fragment (EcoI3 or EcoC) carrying a portion of the T-R DNA (Barker et al., Plant Mol. Biol. (1983) 2:325) including the  
20 mannopine synthase promoter region ( $P_{MAS}$ ) was cloned in a vector designated pVK232. After digestion of pVK232 with EcoRI, EcoI3 was inserted into the EcoRI site of pACYC184 to yield plasmid pCGN14. pCGN14 was digested  
25 with SphI and ClaI (respectively at position 21562 and 20128 of the Barker et al. sequence, supra) to remove the  $P_{MAS}$  region which was inserted into pUC19 (Pharmacia, Inc.) which had been digested with SphI and AccI to yield pCGN40. The  $P_{MAS}$  region includes a ClaI  
30 recognition site internally which is methylated, so as to resist digestion.

pCGN40 was digested with EcoRV and EcoRI where the EcoRV site is in the T-DNA, while the EcoRI site is in the polylinker of pUC19 to provide a  
35 fragment having the  $P_{MAS}$  region. pCGN451 containing the octopine synthase cassette was digested with SmaI and EcoRI and the larger fragment isolated from which

the octopine synthase 5' region had been removed. The EcoRV-EcoRI P<sub>MAS</sub> region was substituted into pCGN451 for the octopine synthase 5' region, where the transcriptional initiation and termination regions were separated by a polylinker to provide pCGN46.

To introduce the plasmid pCGN809 onto the wide host range T-DNA binary vector pCGN587, pCGN809 was transformed into an E. coli polA1 mutant containing pCGN587. The two plasmids pCGN587 and pCGN809 share two regions of homology where recombination can occur: at the ocs 3' regions of the APHII gene and the  $\gamma$ -interferon gene, and at the pUC origins of the two plasmids, recombination at either site results in a plasmid containing both the  $\gamma$ -interferon gene and the APHII gene as intact expression units. In the polA1 mutant, the pCGN809 plasmid, which has a pUC origin of replication, can only survive by integration into the pCGN587 plasmid, which has a wide host range replication system. Since the pCGN809 construct has ampicillin resistance, the chimeric plasmid pCGN810 can be selected. The transformed E. coli host was grown in a nutrient medium containing 100  $\mu$ g/ml ampicillin and surviving clones isolated. Plasmid pCGN810 was then mated into A. tumefaciens C58 (Watson et al., J. Bacteriol. (1974) 123:255) by the method of Ditta et al., Proc. Natl. Acad. Sci. USA (1978) 77:7347. pCGN810 was also mated into A. tumefaciens K61, which contains a disarmed Ti-plasmid, as compared to C58 which has the genetic capability for nopaline expression, and is an armed Ti-plasmid. The transformed Agrobacterium was selected with carbenicillin (50  $\mu$ g/ml) and streptomycin (150  $\mu$ g/ml) at 30°C for two to three successive streak outs. After growth under selection, DNA was isolated and restricted with BamHI and EcoRI and verified by probe hybridization. The cloned pCGN808 was used as the nick-translated probe and confirmed the presence of a

558bp  $\gamma$ -interferon fragment in both the expression vector pCGN46 and the T-DNA plasmid pCGN587 in armed and disarmed Agrobacterium.

Four pCGN810 constructs in both C58 and K61 were cocultivated with Nicotiana tabacum xanthii protoplasts and clumps of transformed cells which were selected in 100  $\mu$ g/ml kanamycin were transferred to solid media containing kanamycin (100  $\mu$ g/ml) for further growth and regeneration into plants.

Tobacco leaf disks were cocultivated with pCGN810 in C58 and K61 Agrobacterium strains and the disks transferred to selective media with kanamycin (100  $\mu$ g/ml) for growth to callus.

Sterile tobacco plants were stem-stab inoculated with A. tumefaciens C58 (pCGN810) resulting in the production of gall tissue which was excised and transferred onto solid media containing kanamycin (100  $\mu$ g/ml). After considerable callus proliferation, most of the gall-induced callus material was harvested for mRNA and protein analysis. A small portion of the gall-callus material started to shoot on the kanamycin medium without hormones. Once transferred to shooting media, the shoot development looked normal for most of the plantlets, though some were teratoma-like. The plantlets were transferred to rooting medium containing kanamycin (100  $\mu$ g/ml) and most of the morphology normal plantlets rooted.

The rest of the gall-induced callus material was pooled and used for preliminary expression analysis for  $\gamma$ -interferon as compared with control callus (plant cells transformed with A. tumefaciens C58(pCGN587)). After extraction of the plant material and immunoprecipitation with antisera against  $\gamma$ -interferon, the extract was electrophoresed on 15% SDS-PAGE, electroblotted to nitrocellulose and blocked with  $\gamma$ -interferon antisera. Radioactively labeled S. aureus protein A was hybridized to the filter, so as to bind



to any immune complexes which had formed. In the transformed plant tissue (A. tumefaciens C58 (pCGN810)), a visible band ran at the same size as the positive control  $\gamma$ -interferon samples.

5           Analysis of the mRNA obtained after an oligo-dT prep of total RNA was performed using nick-translated isolated gene fragments to the Tn5 kanamycin gene (in pCGN587) and the  $\gamma$ -interferon fragment (in pCGN808). The Northern data showed a  
10       message in both pCGN587 and pCGN810 mRNA preps which hybridize to the kanamycin fragment. Only the transformed pCGN810 sample, hybridized to the  $\gamma$ -interferon gene probe showing two mRNA species, each terminating in a different sized noncoding 3'-sequence.

15           Samples were prepared and bioassayed for murine  $\gamma$ -interferon. The samples were prepared as follows: To frozen tobacco tissue from tissue culture (1.6 or 1.0g) ground in liquid nitrogen was added 5 or 2.5ml, respectively, extraction buffer (1M NaCl,  
20       25mg/ml BSA, 10mM DTT, 10mM cysteine, 20mM NaPO<sub>4</sub>, pH 7.4) and grinding continued until the mixture was liquid. The liquid was then centrifuged twice at 16K rpm for 20min, transferring the liquid to a fresh tube after the first separation. The sample was then  
25       diluted to 5ml and 0.5ml control pore glass beads (CPG) (washed and autoclaved) were added and the mixture agitated at 4°C for 4hr. The mixture was then loaded onto a silane treated Pasteur pipette column, the column washed with 3ml 20mM NaPO<sub>4</sub>, 1M NaCl and the  
30       product eluted with 5ml 1M NaCl, 30% ethylene glycol. The material eluted from the column was then dialyzed overnight against 20mM NaPO<sub>4</sub>, 1M NaCl. Dialysis buffer was changed once at 3hr. The dialysed sample was concentrated with a Millipore immersible-CX 10,000  
35       ultrafilter to less than about 1ml and then a series of dilutions with BSA buffer (1M NaCl, 20mM NaPO<sub>4</sub>, pH 7.4, 25mg/ml BSA) performed. Comparison samples were made

from transformed and untransformed tobacco, where no  $\gamma$ -interferon gene was introduced, where the tissue was or was not supplemented with  $\gamma$ -interferon. The following table indicates the results.

5

TABLE

	Tissue <sup>1</sup>	Dilution <sup>2</sup>	$\gamma$ -IFN activity <sup>3</sup>
10	UT A	0 1 2 3	9 9 9 9
15	UT B	0 1 2 3	741 188 15 9
20	T A	0 1 2 3	9 9 9 9
	T B	0 1 2 3	9191 9191 1042 261
25	T- $\gamma$ IFN	0 1 2 3	197 41 9 9

- 30 <sup>1</sup> UT - untransformed tobacco tissue  
 T - tobacco tissue transformed with foreign DNA  
 T- $\gamma$ IFN - tobacco tissue transformed with the murine  $\gamma$ -IFN gene  
 A - no murine  $\gamma$ -IFN added  
 B - 5 $\mu$ g  $\gamma$ -IFN added to 5ml of sample prior to purification treatment

35 <sup>2</sup> 0 = 0.5; 1 = 10<sup>-1</sup>; 2 = 10<sup>-2</sup>; 3 = 10<sup>-3</sup>

<sup>3</sup> Assay procedure is described in Yip *et al.*,  
Proc. Natl. Acad. Sci. USA (1982) 79:1820-1824.

The above results demonstrate that one can produce physiologically active mammalian proteins in plant cells. Stable messages are produced by transcription of the mammalian gene integrated into the plant genome, which message can then be translated into the mammalian product, so as to provide physiologically active proteins. The mammalian proteins can be readily isolated free of deleterious contaminants, such as oncogenic DNA, exotoxins, and the like. In addition, the products can be processed to provide for a mature product having the same or substantially the same structure and composition as the naturally-occurring peptide product.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

20

25

30

35

## WHAT IS CLAIMED IS:

1. A method for producing a mammalian peptide which comprises:

growing plant cells containing an integrated  
5 sequence comprising,

a first expression cassette having in the  
direction of transcription (1) a transcriptional and  
translational initiation region functional in said  
plant cells, (2) a structural gene coding for said  
10 mammalian peptide, and (3) a termination region,

whereby said structural gene is expressed to  
produce said mammalian peptide; and

isolating said mammalian peptide  
substantially free of plant cell components.

15

2. A method according to Claim 1, wherein  
said integrated sequence includes a second expression  
cassette having in the direction of transcription (1) a  
transcriptional and translational initiation region  
20 functional in said plants, (2) a structural gene coding  
for a peptide which allows for selection of plant cells  
expressing said peptide, and (3) a termination region.

3. A method according to Claim 2, wherein  
25 said transcriptional and translational initiation  
region is derived at least in part from a  
transcriptional and translational initiation region of  
a Ti- or Ri-plasmid.

4. A method according to Claim 3, wherein  
30 said transcriptional and translational initiation  
region regulates expression of mannopine synthase,  
octopine synthase or nopaline synthase.

5. A method according to Claim 1, wherein  
35 said transcriptional and translational initiation

region of said first expression cassette regulates expression of a plant gene.

6. A method according to Claim 1, wherein  
5 said integrated sequence includes a boundary region from T-DNA.

7. A method for producing an interferon which comprises:  
10 growing plant cells containing an integrated sequence comprising,  
a first expression cassette having in the direction of transcription (1) a transcriptional and translational initiation region functional in said  
15 plant cells and derived from a region which regulates expression of a T-DNA gene, (2) a structural gene coding for an interferon, and (3) a termination region functional in said plant cells,  
whereby said structural gene is expressed to  
20 produce said interferon, and  
isolating said interferon substantially free of plant cell components.

8. A method according to Claim 7, wherein  
25 said integrated sequence includes a second expression cassette having in the direction of transcription (1) a transcriptional and translational initiation region functional in said plant cells, (2) a structural gene coding for an enzyme which imparts antibiotic  
30 resistance, and (3) a T-DNA boundary.

9. A method according to Claim 8, wherein  
said first expression cassette transcriptional and translational initiation region regulates expression of  
35 the mannopine synthase gene of T-DNA.

10. An expression cassette comprising a DNA sequence having in the direction of transcription a transcriptional and translational initiation region functional in plant cells, a structural gene coding for a mammalian peptide, and a termination region functional in plant cells.

11. An expression cassette according to Claim 10 including joined to said DNA sequence a second expression cassette comprising a second transcriptional and translational initiation region functional in plant cells, a structural gene coding for a peptide providing a phenotypic property capable of selection in plant cells, and a termination region functional in plant cells.

12. An expression cassette according to Claim 11, including a T-DNA boundary.

13. A DNA construct comprising a first expression cassette having in the direction of transcription a transcriptional and translational initiation region regulating the expression of mannopine synthase of T-DNA, a structural gene coding for  $\gamma$ -interferon and a termination region functional in plant cells, a second expression cassette comprising in the direction of transcription a transcriptional and translational initiation region regulatory the expression of octopine synthase of T-DNA, a structural gene coding for an enzyme imparting antibiotic resistance to plant cells, and a termination region functional in plant cells, and a T-DNA boundary.

# INTERNATIONAL SEARCH REPORT

International Application No PCT/US86/1599

<b>I. CLASSIFICATION OF SUBJECT MATTER</b> (If several classification symbols apply, indicate all) <sup>1</sup>		
According to International Patent Classification (IPC) or to both National Classification and IPC IPC(4): C12P 21/00; C12N 15/00; C07H 21/04 U.S. CL: 435/68, 172.3; 536/27		
<b>II. FIELDS SEARCHED</b>		
Minimum Documentation Searched <sup>2</sup>		
Classification System	Classification Symbols	
U.S.	435/68, 172.3; 536/27	
Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched <sup>3</sup>		
COMPUTER SEARCH CAS, BIOSIS 1975 TO PRESENT MAMMAL, PEPTIDE, GENE, INTERFERON, PLANT, AUTHORS		
<b>III. DOCUMENTS CONSIDERED TO BE RELEVANT</b> <sup>14</sup>		
Category <sup>6</sup>	Citation of Document, <sup>15</sup> with indication, where appropriate, of the relevant passages <sup>17</sup>	Relevant to Claim No. <sup>18</sup>
A	US, A, 4,399,216 (AXEL) 16 August, 1983	1, 2, 7, 10, 11
A	US, A, 4,530,901 (WEISSMAN) 23 July, 1985	1, 7, 10
X	EP, A, 0,108,580 (AUSICH) 16 May, 1984	1, 2, 6, 7, 8, 10, 11, 12
Y	A. DEPICKER ET AL., "Plant Cell Transformation by Agrobacterium Plasmids" in <u>Genetic Engineering of Plants-An Agricultural Perspective</u> , 1983, Plenum Press (New York, New York). See pages 144 to 148.	3, 4, 5, 9, 13
Y	<u>Proceedings of the National Academy Sciences USA</u> , Volume 80, issued August, 1983, (Washington, D.C.), R.T. Fraley et al., "Expression of Bacterial Genes in Plant Cells" pages 4803-4807.	3, 4, 5, 9, 13
<p><sup>5</sup> Special categories of cited documents: <sup>16</sup></p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier document but published on or after the international filing date</p> <p>"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p> <p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step</p> <p>"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.</p> <p>"A" document member of the same patent family</p>		
<b>IV. CERTIFICATION</b>		
Date of the Actual Completion of the International Search <sup>1</sup>		Date of Mailing of this International Search Report <sup>1</sup>
October 10, 1986		07 NOV 1986
International Searching Authority <sup>1</sup>		Signature of Authorized Officer <sup>1</sup>
ISA/US		Debra F. Edwards

## INFORMATION CONTINUED FROM THE SECOND SHEET

Name of Applicant: Smith, William Issued September 1, 1986  
1986, Amsterdam, Holland, NL  
 Title of Invention: General Method for the Transfer of Cloned Genes to Plant Cells, pages 315-330.

1,10

## V. OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. ☐ Claim numbers because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claim numbers because they relate to parts of the international application that did not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out (2), specifically:

## VI. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This International Searching Authority found multiple inventions in this international application as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. ☐ As all searchable claims could be searched without effecting an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest:

☐ The additional search fees were accompanied by applicant's protest.

☐ No protest accompanied the payment of additional search fees.